

Wie wirkt sich die Verwendung von KI auf den Lernerfolg von Software-Engineering-Studenten aus?

Frank Mayer

Software Engineering

Hochschule Heilbronn

Heilbronn, Deutschland

fmay2@stud.hs-heilbronn.de

Abstract—In jüngster Zeit gewinnen KI-gestützte Tools zunehmend an Bedeutung im Bildungsbereich, wobei ihr Einsatzpotenzial für das Lernen von Software-Engineering-Studierenden bisher unzureichend erforscht ist. Diese Arbeit untersucht, ob und unter welchen Bedingungen KI-basierte Tools den Lernerfolg von Studierenden in Software-Engineering-Kursen positiv beeinflussen können.

Der Untersuchungsrahmen ist die Lehrveranstaltung “Interaktive Programme” im ersten Semester des Studiengangs Software Engineering an der Hochschule Heilbronn. In einem quasi-experimentellen Design werden zwei Gruppen von Studierenden miteinander verglichen: eine Kontrollgruppe ohne Einsatz von KI-Tools und eine Experimentaltgruppe, die spezifisch auf KI-gestützte Tutorien zurückgreifen soll. Neben der Bewertung der Aufgabenlösungen werden qualitative Beobachtungen und Selbsteinschätzungen der Studierenden genutzt, um den Einfluss der KI-Tools auf Lernerfolg, Motivation und Selbsteinschätzung zu analysieren.

Die Ergebnisse dieser Studie sollen Hinweise darauf geben, wie KI-Tools im Kontext der Hochschullehre sinnvoll integriert werden können, um den Lernerfolg langfristig zu fördern, ohne das eigenständige Lernen der Studierenden zu beeinträchtigen.

Index Terms—Artificial Intelligence in Education, Generative AI, Higher Education, Software Engineering Education, Learning Outcomes, Self-assessment, Tutoring Systems, Quasi-experimental Design, Qualitative Observation, Educational Technology, Student Motivation

I. EINFÜHRUNG

KI-basierte Werkzeuge, die speziell für den Einsatz im Bildungsbereich geeignet sind, sind ein vergleichsweise neues Phänomen. Ziel dieser Arbeit ist es, zu untersuchen, ob und unter welchen Bedingungen KI-basierte Tools den Lernerfolg von Studierenden im Bereich Software Engineering fördern können.

Das Bedürfnis, dieses Thema zu erforschen, ergibt sich aus meiner eigenen Erfahrung als Software-Engineering-Studierender. Ich beobachte bei Kommilitonen, dass Tools wie ChatGPT verwendet werden, um Aufgaben zu erleichtern. Dabei treten jedoch Bedenken auf: Häufig wird die Qualität der generierten Informationen nicht hinterfragt, und es mangelt an einem Bewusstsein für den eigenen Lernprozess und die eigene Kompetenzentwicklung.

Daraus ergibt sich die Frage, ob diese Problematik auf die Eigenschaften der Tools selbst zurückzuführen ist oder auf den Umgang der Studierenden mit diesen Werkzeugen.

Das Ziel dieser Arbeit ist es daher, herauszufinden, wie der Einsatz von KI-Tools im Software-Engineering-Studium gestaltet werden sollte, um eine optimale Ausbildung der Studierenden zu gewährleisten. Die Frage nach der generellen Eignung von KI für die berufliche Praxis von Softwareentwicklern wird dabei nicht thematisiert.

Einige Studien deuten darauf hin, dass der Einsatz personalisierter KI-Tutoren, die auf Lernstrategien wie Abstands-lernen und gezielte Wiederholung setzen, die Lernleistungen signifikant steigern kann. Eine Untersuchung an der UniDistance Suisse ergab, dass Studierende, die einen solchen KI-Tutor nutzten, eine Verbesserung der Durchschnittsnoten um bis zu 15 Perzentilpunkte erzielten. [1]

Eine weitere Studie zeigt, dass ein individualisiertes KI-Tutorsystem, das auf drei Developmental Learning Networks (DLNs) basiert und als mobile App zur Sprachvermittlung (Koreanisch) für Kinder implementiert ist, den Lernfortschritt signifikant fördert. Dabei wurde der Lernstand kontinuierlich erfasst und die App an die individuellen Präferenzen der Kinder angepasst. [2]

II. FORSCHUNGSFRAGEN

A. Hauptfrage

Inwiefern beeinflusst der Einsatz KI-basierter Tools den Lernerfolg von Studierenden in Hochschulkursen im Bereich Software Engineering?

B. Unterfragen

- 1) Wie verändert sich im Verlauf des Studiums die Selbsteinschätzung der Studierenden in Bezug auf ihren Wissensstand?
- 2) Wie unterscheidet sich der Umgang der Studierenden mit KI-Tools im Vergleich zur Interaktion mit menschlichen Tutoren?

- 3) In welchem Verhältnis steht die durch KI-Tools geförderte Selbsteinschätzung der Studierenden zu deren tatsächlich gemessenem Lernerfolg?

III. METHODIK

A. *Setting*

Die Untersuchung wird im Rahmen der Vorlesung „Interaktive Programme“ an der Hochschule Heilbronn durchgeführt. Diese Vorlesung ist Teil des ersten Semesters des Bachelorstudiengangs Software Engineering und vermittelt grundlegende Programmierkenntnisse. Im Rahmen der Studie werden die Studierenden gezielt sowohl mit als auch ohne KI-basierte Tools lernen.

Zusätzlich zur Vorlesung wird ein Tutorium angeboten, das als ergänzende Veranstaltung dient. Die Teilnahme am Tutorium ist freiwillig, wird jedoch von der Lehrperson besonders den Studierenden empfohlen, die Schwierigkeiten mit den Lerninhalten haben.

B. *Forschungsdesign*

Für diese Studie wird ein quasi-experimentelles Design verwendet, bei dem zwei Gruppen von Studierenden untersucht werden: eine Kontrollgruppe (KG) und eine Experimentalgruppe (EG).

Beide Gruppen werden aus den Teilnehmenden der Vorlesung „Interaktive Programme“ zusammengestellt.

C. *Gruppen*

Die Studierenden werden in zwei Gruppen aufgeteilt:

- **KG:** Nicht-KI-Gruppe
Die Studierenden in der Kontrollgruppe werden dazu angehalten, keine KI-Tools zu verwenden. Sie erhalten ein traditionelles Tutorium, in dem ich sie bei der Problemlösung unterstütze und ihnen Erklärungen zur Aufgabenstellung biete, ohne konkrete Lösungen zu geben. Stattdessen erhalten sie kleine Hinweise, die sie in Richtung einer möglichen Lösung führen, sodass sie die Lösung eigenständig erarbeiten können.
Da die Studierenden im ersten Semester keine umfangreichen Vorkenntnisse mitbringen, werde ich sie über KI-basierte Tools informieren und auf Tools hinweisen, die aufgrund ihrer Funktionalität nicht verwendet werden sollten. Ich kann jedoch nicht vollständig verhindern, dass Studierende außerhalb des Tutoriums dennoch auf KI-Tools zurückgreifen.
- **EG:** KI-Tutor-Gruppe
Die Studierenden in der Experimentalgruppe werden dazu ermutigt, KI-gestützte Tools zur Unterstützung ihres Lernprozesses zu nutzen. Konkret sollen Chatbots wie Claude für Verständnisfragen herangezogen werden. Der Chatbot wird so konfiguriert, dass er sich wie ein Tutor verhält, der die Studierenden durch den Lösungsweg führt, ohne konkrete Lösungen zu liefern. Um sicherzustellen, dass die Studierenden die KI-Tools

effizient einsetzen können, erhalten sie von mir eine Einführung und Unterstützung bei deren Anwendung. Während Tests und Klausuren ist die Nutzung von KI-Tools untersagt. Diese Einschränkung soll auch dazu beitragen, die Ergebnisse der Studie „Generative AI Can Harm Learning“ [3] zu überprüfen, die zeigt, dass die Nutzung von KI-Tools kurzfristig zu Lernfortschritten führen kann, jedoch im Anschluss, ohne KI-Unterstützung, zu Leistungseinbußen im Vergleich zur Kontrollgruppe.

Die Aufgaben stammen regulär aus der Vorlesung und werden unverändert übernommen. Die Bewertung erfolgt unabhängig von der Gruppenzugehörigkeit.

Die Teilnahme an der Studie steht allen Studierenden des Studiengangs Software Engineering, die die Vorlesung „Interaktive Programme“ besuchen, offen und erfolgt auf freiwilliger Basis. Es werden keine Anreize zur Teilnahme angeboten.

D. *Gruppenzuteilung*

Zu Beginn des Semesters werden alle Studierenden gefragt, ob sie an der Studie teilnehmen möchten. Sie werden über das Ziel der Studie sowie die Zusammensetzung der Gruppen informiert. Die Zuteilung zu den Gruppen erfolgt unter Berücksichtigung der bisherigen Leistungsbewertung, sodass die Gruppen möglichst gleichmäßig verteilt sind.

E. *Verwendete Tools*

Die Studierenden der Experimentalgruppe (EG) können frei wählen, welche KI-Tools sie verwenden möchten. Ich werde jedoch einige Tools empfehlen, die sich als geeignet erwiesen haben, und eine Einführung in deren Nutzung geben. Die empfohlenen Tools sind:

- **Claude**
Ein auf Claude 3.5 Sonnet basierender Chatbot.
- **Microsoft Copilot**
Ein auf dem GPT-4-Modell von OpenAI basierender Chatbot. Der Zugriff ist über den Hochschulaccount kostenlos möglich.
- **ChatGPT**
Ein Chatbot basierend auf OpenAI's GPT-4o und GPT-4o mini Modellen. Der Zugang ist ebenfalls über den Hochschulaccount kostenlos verfügbar.

Auf Grundlage meiner persönlichen Erfahrungen eignet sich Claude besonders gut als Tutor, da dieses LLM versucht, die Ursache eines Problems zu identifizieren, bevor es Lösungsvorschläge anbietet. GPT hingegen generiert sofort Lösungsvorschläge, die oft nicht auf das eigentliche Problem eingehen, wenn es nicht zuvor klar vom Nutzer formuliert wurde.

F. Datenerhebung

a) Messung:

- 1) Der Lernerfolg wird anhand der Bewertung der Aufgaben gemessen. Dabei wird bewusst auf Selbsteinschätzungen der Studierenden verzichtet, da diese oft unzuverlässig sind. Die Bewertung erfolgt durch uns Tutoren. Der Professor stellt eine Musterlösung zur Verfügung, und wir besprechen gemeinsame Bewertungsstandards. Jede Teilaufgabe wird vollständig von einem Tutor bewertet, um die Vergleichbarkeit zu gewährleisten.
- 2) Der Umgang der Studierenden mit dem Tutor bzw. den KI-Tools wird durch gezielte Beobachtungen erhoben. Ich beobachte die Studierenden, die mit KI-Tools arbeiten, während mein Kollege die Studierenden beobachtet, die im Tutorium Unterstützung erhalten. Diese Beobachtungen werden in einem Protokoll festgehalten, das folgende Fragen enthält:
 - Wie präzise formulieren die Studierenden ihre Fragen?
 - Wie gut verstehen die Studierenden die Antworten des Tutors/der KI?
 - Fragen die Studierenden nach, wenn sie die Antwort nicht verstanden haben? (sichtbares Nichtverstehen: z. B. Zögern bei der Umsetzung, Nachfrage bei Kommilitonen)
- 3) Die Selbsteinschätzung der Studierenden wird mithilfe eines Fragebogens erhoben, der direkt die subjektive Wahrnehmung des eigenen Lernfortschritts, der Leistung, der Motivation und des Wissensstands erfasst. Der Fragebogen wird nach jedem Tutorium ausgefüllt und basiert auf einer Skala von 1 bis 5, wobei 1 für „stimme nicht zu“ und 5 für „stimme voll zu“ steht. Wenn ein Studierender am Tutorium nicht teilnimmt, entfällt auch die Selbsteinschätzung. Der Fragebogen enthält folgende Aussagen:
 - Die Aufgabenstellung war für mich ohne Hilfe verständlich (1-5).
 - Die Hilfestellung des Tutors/der KI war für mich hilfreich (1-5).
 - Ich konnte die Aufgabe ohne Hilfe lösen (1-5).
 - Die Unterstützung des Tutors/der KI hat meine Lösungsansätze beeinflusst (1-5).
 - Den Antworten des Tutors/der KI habe ich vertraut (1-5).
 - Mein Wissensstand hat sich durch die Aufgabe verbessert (1-5).
 - Ich fühle mich sicher, eine ähnliche Aufgabe ohne Hilfe lösen zu können (1-5).
 - Sonstige Anmerkungen (Freitextfeld).Der tatsächliche Lernerfolg wird durch die Bewertung der abgegebenen Lösungen ermittelt.

b) Auswertung:

- 1) Bewertung der Aufgaben als Maß für den Lernerfolg: Alle Aufgaben werden von meinem Kollegen und mir bewertet. Jede Teilaufgabe wird dabei vollständig von einem Tutor bewertet, um eine einheitliche Bewertung zu gewährleisten. Die erzielten Punkte sowie ihre Entwicklung im Verlauf des Semesters werden sowohl auf individueller Basis als auch als Median und Durchschnitt innerhalb jeder Gruppe betrachtet.
- 2) Beobachtung des Umgangs mit dem Tutor und den KI-Tools: Die in den Protokollen festgehaltenen Beobachtungen werden qualitativ ausgewertet, wobei besondere Auffälligkeiten und Unterschiede zwischen den Gruppen erfasst werden. Eine quantitative Auswertung ist hierbei nicht vorgesehen; stattdessen sollen aus den qualitativen Beobachtungen relevante Unterschiede und Gemeinsamkeiten abgeleitet werden.
- 3) Selbsteinschätzung der Studierenden: Die Selbsteinschätzung wird anhand des Fragebogens erhoben und ebenfalls qualitativ ausgewertet. Besonderes Augenmerk liegt auf Unterschieden zwischen den Gruppen, insbesondere auf möglichen Überschätzungen in der Experimentalgruppe (EG). Diese Beobachtung basiert auf meinen bisherigen Erfahrungen und soll empirisch überprüft werden.

G. Aufgaben

wird ergänzt wenn bekannt

REFERENCES

- [1] A. Baillifard, M. Gabella, P. B. Lavenex, and C. S. Martarelli, "Effective learning with a personal AI tutor: A case study," *Education and Information Technologies*, pp. 1–16, 2024.
- [2] W.-H. Kim and J.-H. Kim, "Individualized AI tutor based on developmental learning networks," *IEEE Access*, vol. 8, 2020.
- [3] H. Bastani, O. Bastani, A. Sungu, H. Ge, Ö. Kabakçı, and R. Mariman, "Generative AI Can Harm Learning," *Available at SSRN 4895486*, 2024.